An Roinn Iompair
Department of Transport

Department of Transport Data Strategy

2024-2030

Table of Contents

Introduction **4**
History of Data in the Department **5**
Principles **7**
Data Maturity **9**
Business Opportunities **11**
    Current Business Needs 12
    Strategic Priorities 14
    Future Opportunities 15
    Ambitious Use Cases 20
Data Analytics Platform Architecture **21**
Governance **24**
    Data Quality 25
    Unique Identifiers 26
    Security and Data Protection 26
    Data Sharing 26
    Data Discoverability and Metadata 27
Operating Model **29**
    Federated Model for Analytics and BI 29
    DAAD Overall Responsibilities 29
    Data Quality 29
    Data Entering the Department 29
    Data Quality within the Analytics Platform 30
    Data Security 30
    Data Sharing 30
    NVDF Data 30
    Technology 31
Data and the Department's Agencies **32**
DAAD and other Divisions in the Department **33**
Linkages with CSO and Open Data **34**
Roadmap **36**

3

Introduction

In January 2023 the Department of Transport created a new Data and Analytics Division (DAAD). The division was created to ensure optimal use of the Department's data holdings, those of its agencies and other relevant data to inform and evaluate policy. The division is also responsible for ensuring the Department contributes to the National Data Infrastructure (NDI) and to public sector open data initiatives.

The purpose of this data strategy is to layout the intended strategic insights the Department plans to gain from data and to describe the operating model, governance structures, technology, skills, and organisational change required to achieve these.

This document begins with a brief history of data in the Department of Transport, including previous data strategies and current gaps. This data strategy is founded on a number of clear principles which are set out at the beginning. The current data maturity level in the Department is outlined next, followed by the strategic insights to be gained from data. These are categorised and range from known current business requirements, through to strategic and future needs and then ambitious use cases. The categories of insight will drive the evolution of a data architecture which is described next. The operating model is presented, setting out a vision where the Data and Analytics Division becomes a 'Centre of Excellence' for the Department as part of a federated model. DAAD will be responsible for enabling the use of data throughout the department, including making cleansed, standardised, well described data available with the correct security, and identity management processes in place and by providing the correct toolset. DAAD will help upskill divisions in analytic methods and will create and deploy models into production. The operating model will outline ownership for the various components of the data function such as technology, data governance, data quality and management, analysis, and modelling. The document goes on to describe the planned role of the agencies and the interdependencies between the Department, other Public Service Bodies (PSBs) and the Central Statistics Office (CSO).

Finally, this strategy presents a roadmap describing how and when the Department plans to achieve the goals set out.

History of Data in the Department

Work has previously been undertaken in assessing data availability and use within the Department. The CSO published a number of ‘Statistical Potential of Administrative Records’ (SPAR) reports and the 2003 report analysed data collection methods in six departments including Transport. The report made a number of recommendations including, but not limited to:

- Developing a national transport or mobility survey.
- Establishing a data or statistics section within the Department.
- Improving definition of performance and strategic goals.
- Improving measurement of congestion.
- Better sharing of data between agencies.

Following the SPAR Report the Department published the ‘Department of Transport Data and Statistics Strategy 2009-2011’. Significant outcomes from this strategy included the creation of a National Travel Survey for the first time, the creation of a Transport Statistics Liaison Group and the closing of some data gaps identified.

In 2017 the Department of Transport, Tourism and Sport published a Data Strategy for the period 2017-2019. The planned actions in this strategy included allocating dedicated statistics and data resources, improving data quality and management, developing a data hub to allow access to departmental data and improving the availability and use of measurable indicators that enable direct assessment of output/outcomes against objectives. The key division in the department responsible for many of these actions at the time was the Economic and Financial Evaluation Unit (EFEU), now the Strategic Research and Analysis Division (SRAD).

Currently SRAD provides a strategic analytical and research resource to support the Department and its agencies to ensure policy development, monitoring, and evaluation is evidence based and objectives led. Key objectives of the division include:

- Compliance with Economic Appraisal Requirements within the Public Spending Code.

- Strategic Research (e.g., spending reviews).

- Analytical Support (reviewing analysis underpinning a proposal or carrying out background research to inform the development of a new policy).

- Research Coordination.

- Lead liaison to the International Transport Forum.

The creation of DAAD in 2023 provides the dedicated focus required to advance some of the key data actions that have been highlighted for a number of years while allowing SRAD to continue the vital role it plays.

Principles

This data strategy is founded on six clear principles which will guide decision making as the data ecosystem matures in the Department.

1. Promote and Enable a Data Driven Culture

> Data are valued and used throughout the Department & agencies to provide clear insight for transport policy and for provision of services to citizens.

2. Data are Linkable

> The data assets of the Department are well populated with the appropriate unique identifiers to allow linking to other public data sets.

3. Data are Discoverable, Accessible and Shared Safely

> Data are easily discovered via a searchable catalogue. Data are accessible to the consumer in an appropriate form. Data described by high quality metadata. Data sharing will be encouraged and will comply with the appropriate legislation.

4. Data are Governed and Protected

> Data are collected and processed according to legislation. Data are of high quality, cleansed and standardised. Data will be secured with clear identity management procedures. Specific procedures and protections will be in place for personal data. Appropriate retention and archiving policies apply to stored data.

5. Data are Centralised

> Data for analytics will be stored in one central repository.

6. Continuous Engagement

> User needs will continually be assessed, the use of data and data skills will be advocated for.

Data Maturity

Currently the Department has varying levels of capability regarding data and analytics. Some Divisions have high-level research questions but lack staff with analytics skills to derive insights from available data sets. Some areas need information that they currently receive in raw form to be visualised in a dashboard for ease of interpretation and to clearly see trends. Other groups have analytic capabilities but need access to new sources of data and need the current data to be cleansed, standardised, and accessible.

In accordance with the Public Service Data Strategy, the Department's Data and Analytics Division will have overall responsibility for the safe, transparent, and optimised use of data. The division will have the following primary responsibilities that the Public Service Data Strategy sets out as well as others outlined in the operating model:

a. Capability assessment in relation to data.
b. Creation and implementation of a data strategy.
c. Alignment of Department with the Public Service Data Strategy.
d. Work with the Department's Data Protection Officer to ensure compliance with data management governance processes, policies, standards, and guidelines.
e. Leverage data as an asset enabling the organisation to meet its objectives more efficiently and effectively.
f. Act as one point of contact in relation to data management for the organisation.

The desired future state is one where the Data and Analytics Division is a 'Centre of Excellence' for data and analytics in the organisation as part of a federated model as illustrated in Figure 1. DAAD will be responsible for enabling the use of data throughout the Department, including making cleansed, standardised, well described data available with the correct security, and identity management processes in place and by providing the correct toolset. DAAD will help upskill divisions in analytic methods and will create and deploy models into production. This will include machine learning and advanced statistical models over time. DAAD will also play a data stewardship and co-ordination role with our agencies and will have overall responsibility for managing data sharing agreements.

The Department will provide a central data catalogue that is searchable and will have high quality metadata. This will enable all departmental staff to search for available datasets and understand what they contain.

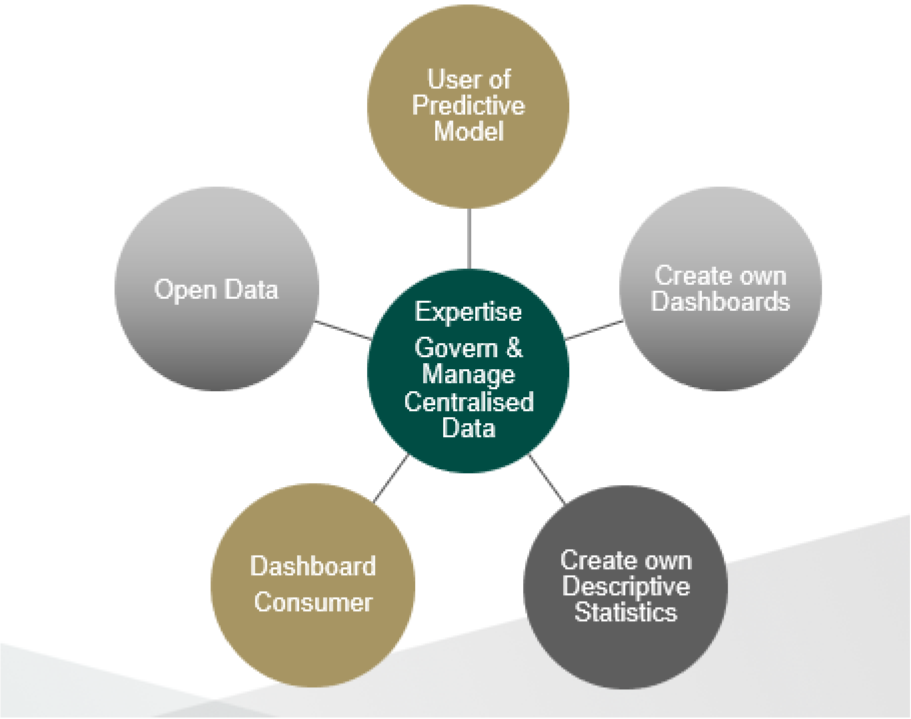

User of Predictive Model
Open Data
Create own Dashboards
Expertise Govern & Manage Centralised Data
Dashboard Consumer
Create own Descriptive Statistics

Figure 1 Centre of Excellence Model

Business Opportunities

This data strategy is focused on the insights to be gained for the Department to inform and enable policy and on the uses of data to reform, modernise and integrate transport services from a citizen perspective. These insights can be grouped into the following categories:

- Current business needs.
- Strategic priorities.
- Future opportunities.
- Ambitious uses of data and analytics.

Each category will drive the evolution of the architecture of an analytics platform. This platform needs to be available in a suitably short timescale to enable the current business needs. It should not be overly complex initially but should have the potential to scale easily to support the needs of the above categories.

In parallel with work on evolving insights as described below, DAAD has created a 'Statistical Spotlight Series' which replaced the annual Transport Trends publication. This new approach offers a more effective way to continuously engage with divisions across the Department and curate new data use cases.

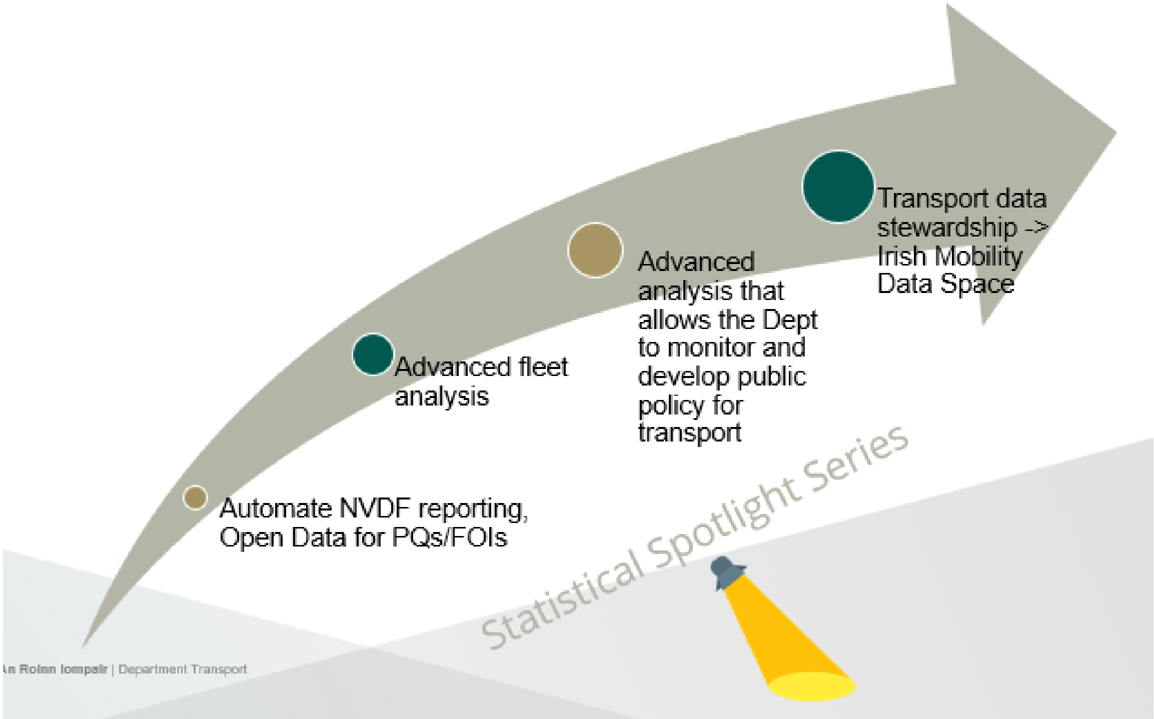

Automate NVDF reporting, Open Data for PQs/FOIs
Advanced fleet analysis
Advanced analysis that allows the Dept to monitor and develop public policy for transport
Transport data stewardship -> Irish Mobility Data Space
Statistical Spotlight Series
An Roinn Iompair | Department Transport

Figure 2 Evolution of Insights

Current Business Needs

The National Vehicle and Driver File (NVDF) is the largest single data holding of the Department. It is made up of two operational databases, the driver file and the vehicle file. The driver file contains information on driver license holders while the vehicle file contains information on all vehicles in the state including owner details, vehicle details and motor tax information. The vehicle file is used for the operation of motor tax which collects close to €1 billion for the state each year.

The NVDF is currently used for analytic purposes by the Department. Its data are used to produce reports to support the smooth operation of the Driver and Vehicle Computer Services Division (DVCSD), including online motor tax. It is also used to answer PQs and FOI requests, mostly pertaining to motor tax, penalty points, driver disqualifications, driver licenses held and breakdowns of vehicles under current taxation. Extracts from the NVDF are also sent periodically to the CSO, published as open data, and transmitted to Eurostat under regulation.

Every year a driver bulletin is produced giving detailed breakdowns of vehicles licenced in Ireland.

This analytic work is currently consuming more time and effort than it should. Data from the live operational system are copied every night to a 'batch server' and copied every month to a 'stats database'. Some of the analytics are produced from the batch server via Tableau and some from the stats database. A lack of standardisation, data cleanse procedures and repeatable scripts are causing consistency issues and making the production of reports overly burdensome.

The most pressing current business need is to streamline this work to increase quality and efficiency. Achieving this requires an analytics platform, a repository where data can be loaded from the operational NVDF system in a repeatable, scheduled manner and is cleansed and standardised and arranged appropriately. This analytics platform should then have the correct toolset for analytics and scripts should be available to facilitate the automation of much of the work described above. Cleansing and standardising the data will most likely require an NVDF data business owner to engage with organisations that provide the source data for the NVDF, e.g., Revenue provide data from VRT. This aspect is discussed further in the Operating Model chapter.

Table 1 - Current Business Needs

Category	**Description**	**Enablers**
Current Business Needs	• Management Information Dashboard for DVCSD • Actionable insights for smooth operation of DVCSD • Streamlined answering of PQs and FOI Requests including through open data. • Driver Bulletin as open data	Cleaned, standardised analytics copy of NVDF

Strategic Priorities

Once the technology and systems are in place to efficiently service the current business needs to a high level of quality, other strategic business use cases can be enabled. Increasing Eircode coverage on the operational NVDF and hence in the copy on the analytics platform will have two clear benefits. Firstly, spatial accuracy will increase, facilitating small area analysis of the vehicle population. This enables insights such as on EV penetration and can inform planning for future electrical supply requirements for home car chargers as well as informing decisions on public charging infrastructure. Secondly, the vehicle file is a rich source of Eircodes due to the high prevalence of car ownership and the frequency of updates to the database (motor tax paid annually or more frequently). An NVDF with well populated Eircodes will contribute to the National Data Infrastructure (NDI). The NDI allows for data from different public service bodies (PSBs) to be linked through unique identifiers. Eircode is the unique identifier for location.

Furthermore, by linking the vehicle and driver file via the PPSN (another key NDI unique identifier) or driver number, richer insights are feasible. For example, the NVDF could be linked to CSO/Revenue income data to help understand the socio-economic profile of EV purchases to inform future grant policies, or to understand how EVs are used as part of multiple car households, or to inform policy to support an equitable transition to EVs.

Table 2 - Strategic Priorities

Category	**Description**	**Enablers**
Strategic Priorities	More nuanced car fleet model modelling for Climate Action Plan and other objectives	Cleaned, standardised analytics copy of NVDF with good Eircode coverage and extra vehicle details.
	Small area statistics on EVs to inform domestic power requirements charging infrastructure promotion of EVs	

	Socio Economic characteristics of EV drivers to inform policy	Linking vehicle to driver via PPSN & further linking with CSO/revenue data

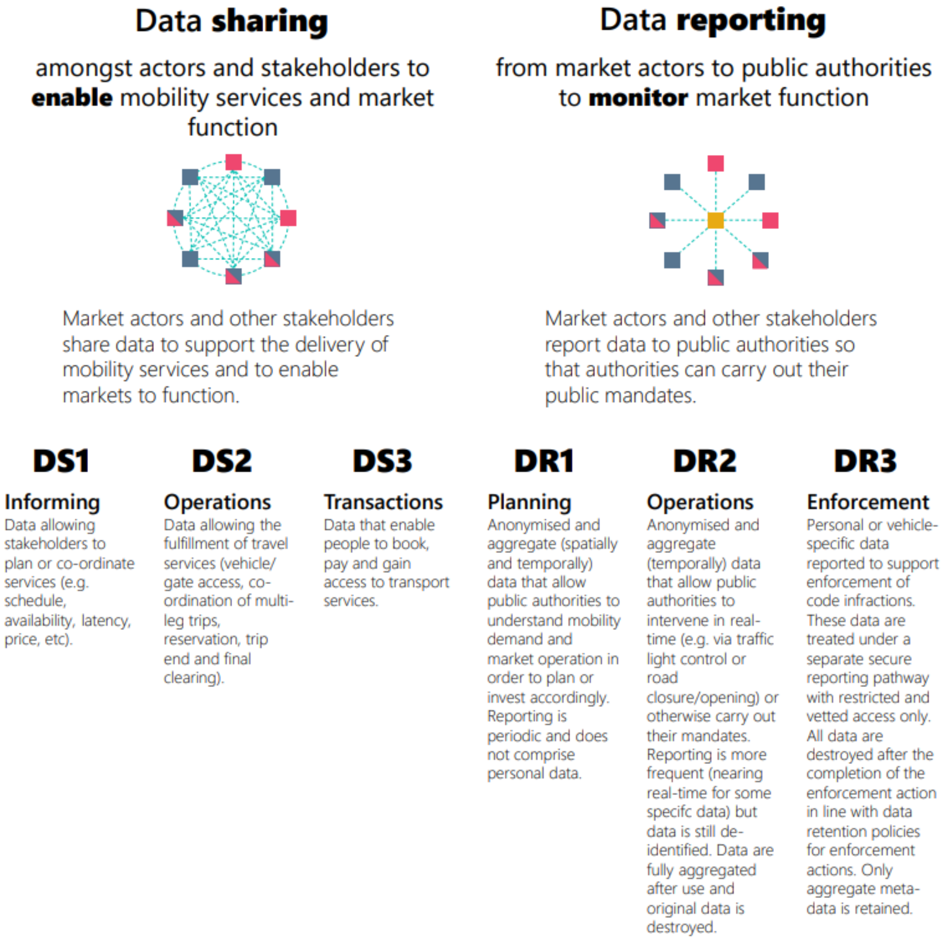
Future Opportunities

Beyond the opportunities afforded by refreshing and streamlining the NVDF, many of the next use cases are enabled by data sharing between the Department and its agencies. This will in turn require further development of the data analytics platform and will also require suitable data sharing agreements to be in place. Some of these future opportunities also involve the Department working collaboratively with other organisations such as CSO.

At this point it is worth referring to a report by the ITF on good practices for reporting mobility data.

Reporting Mobility Data: Good Governance Principles and Practices | ITF (itf-oecd.org)

The report explains *that "mobility data-governance frameworks are comprised of two pillars – data sharing and data reporting. Data sharing refers to data shared among market actors and other stakeholders, which enables the delivery of mobility and other services, and which supports the functioning of transport markets. Data reporting refers to data provided by stakeholders and market actors to public authorities that enables the latter to monitor, guide or intervene to enact public policy."* See Figure 3 Data Sharing vs Data Reporting



Data sharing

amongst actors and stakeholders to **enable** mobility services and market function

Market actors and other stakeholders share data to support the delivery of mobility services and to enable markets to function.

Data reporting

from market actors to public authorities to **monitor** market function

Market actors and other stakeholders report data to public authorities so that authorities can carry out their public mandates.

DS1	DS2	DS3	DR1	DR2	DR3
Informing	**Operations**	**Transactions**	**Planning**	**Operations**	**Enforcement**
Data allowing stakeholders to plan or co-ordinate services (e.g. schedule, availability, latency, price, etc).	Data allowing the fulfillment of travel services (vehicle/ gate access, co-ordination of multi-leg trips, reservation, trip end and final clearing).	Data that enable people to book, pay and gain access to transport services.	Anonymised and aggregate (spatially and temporally) data that allow public authorities to understand mobility demand and market operation in order to plan or invest accordingly. Reporting is periodic and does not comprise personal data.	Anonymised and aggregate (temporally) data that allow public authorities to intervene in real-time (e.g. via traffic light control or road closure/opening) or otherwise carry out their mandates. Reporting is more frequent (nearing real-time for some specifc data) but data is still de-identified. Data are fully aggregated after use and original data is destroyed.	Personal or vehicle-specific data reported to support enforcement of code infractions. These data are treated under a separate secure reporting pathway with restricted and vetted access only. All data are destroyed after the completion of the enforcement action in line with data retention policies for enforcement actions. Only aggregate meta-data is retained.

Figure 3 Data Sharing vs Data Reporting

The Department will have two distinct roles here in respect of each of these two pillars. In terms of the 'Data Sharing' pillar, the Department will have a stewardship role. This will involve ensuring that our agencies collect the correct data in the appropriate formats and with an appropriate quality level. This has two benefits, firstly it improves the agencies' ability to use the data for their operational purposes, but also when the aggregated data is taken back into the Department for analytic purposes it is much more valuable and useable.

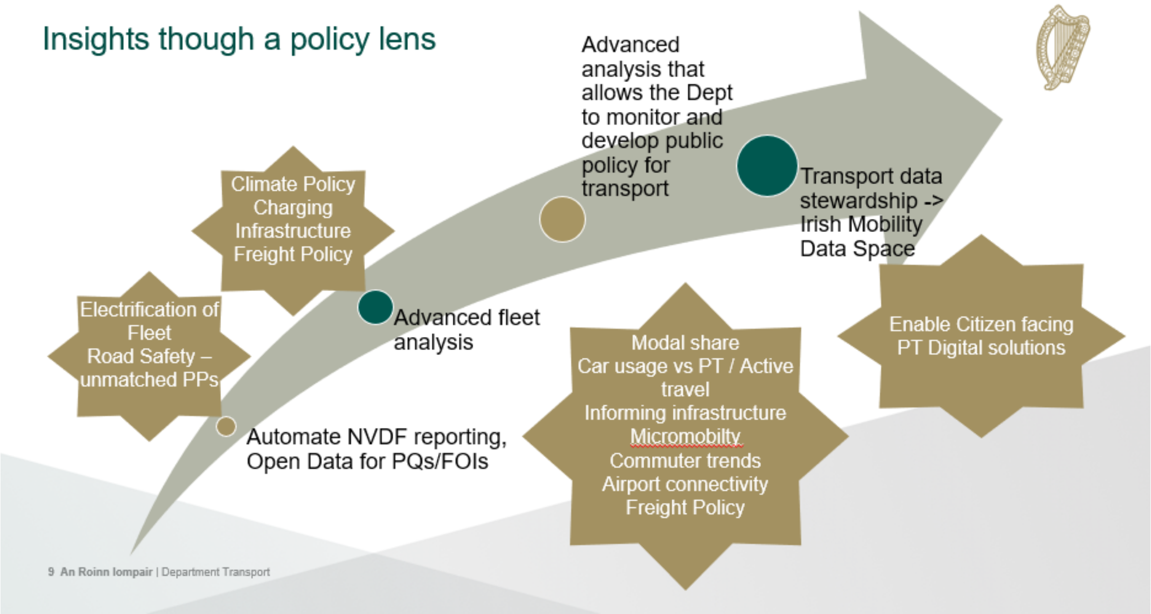
The insights that fall into the "future opportunities" category relate to data reporting that enables the department to monitor and develop public policy for transport (DR1 Planning above).

"Fundamentally, "planning data" are not principally concerned with where individuals

travel but with how communities function (Webb, 2020b). Understanding the latter is more challenging than the former but only requires aggregate insight into mobility patterns. ITF (2021a) addresses many of the issues that must be considered with respect to the use and usefulness of data – especially "big data"– for modelling and planning purposes.
Planning data may enable the creation of origin-destination matrices, which provide insights into how, when and where travel takes place. This information helps governments deliver policies and infrastructure, for example, by providing insights on where new cycling or public transport infrastructure may be needed, where parking may be under- or over-supplied, and where road expansion or re-allocation may be helpful. Alternatively, planning-supportive data may include information on accessibility and mobility service levels. Data reported by mobility operators and other stakeholders may also be used for planning government services or actions outside of the field of transport, for example, in assessing coverage of health or education facilities or adequacy of green space provisions."

Enabling these types of insights will require sharing aggregated data from the NTA related to public transport, cycling, taxi journeys etc, and data from TII on the road network including traffic count data. It may also require data sharing with the RSA. It is not envisaged that raw data (e.g., from transport operators) will be shared with the Department but rather aggregated data. Sharing transport data in this way will allow the Department to combine vehicle, driver, public transport, active travel and road data. This will allow us to gain insights into how car ownership and usage is related to public transport services.

These insights can then be further enhanced in combination with data on housing and population. This would enable a detailed analysis linking where people live and will live with where they work and study and hence help deliver insights on where new active or public transport infrastructure or new road infrastructure might be required.



Insights though a policy lens

Electrification of Fleet
Road Safety – unmatched PPs

Climate Policy
Charging Infrastructure
Freight Policy

Automate NVDF reporting, Open Data for PQs/FOIs

Advanced fleet analysis

Advanced analysis that allows the Dept to monitor and develop public policy for transport

Transport data stewardship -> Irish Mobility Data Space

Modal share
Car usage vs PT / Active travel
Informing infrastructure
Micromobilty
Commuter trends
Airport connectivity
Freight Policy

Enable Citizen facing PT Digital solutions

Figure 4 Insight through a policy lens

Table 3 - Future Opportunities

Category	**Description**	**Enablers**
Future Opportunities	Trends in those using public transport, active travel and micromobility	Data Sharing with NTA
	Airport carbon footprint, volume of cars to/from	Data Sharing with DAA
	Insights into behavioural change towards active travel	Link DoT survey results with NTS and NHTS results
	Analysis of Modal Share	Acquiring smartphone app data Collaborating with NTA and TII to ensure data capture in a consistent way for cycling, shared micro mobility and EV charging
	Deliver insights on where new active or public transport infrastructure or new road infrastructure might be required. Insights into how car ownership and usage is related to public transport services.	Combine vehicle, driver, public transport, and road data. DoT, NTA, TII, RSA data sharing. Then further linked to Census and planning permission data
	Insights on concentration of flows of HGVs Insights on location of distribution hubs on periphery of cities	Collaboration with CSO geospatial expertise & TII traffic count data and TII modelling work
	Demographic profiling of HGV drivers	Collaboration with CSO

	Insights of best location for public EV charging point	EV ownership, population data, deprivation index etc

Ambitious Use Cases

The Department has ambitious plans to reform, modernise and integrate transport services from a citizen perspective. To address this, the Department has created a Digital Hub bringing together its senior leaders in Digital, Technology, Data, Cyber Security and Project Management. Better use of data from within the Department and its agencies will be key to achieving this.

A key enabler for these plans would be a data and services hub which would make relevant data from mobility providers available allowing organisations (public and private) to develop consumer-facing services. Such a hub should create the market conditions for a dynamic range of mobility services.

To orchestrate the creation of such a hub, the Department should liaise with the relevant agencies and act as a data steward to ensure that the correct data are collected, and that standardisation is applied consistently. Data quality metrics should also be specified including accuracy, completeness, consistency, timeliness, uniqueness, validity.

International examples of such mobility data spaces include https://mobility-dataspace.eu/.

Most of the literature on mobility data spaces recommends a decentralised approach to data storage with a centralised catalogue and a clear governance framework to ensure interoperability. Data controllers would decide who can have access to their data.

Data Analytics Platform Architecture

To best govern, manage, protect and share the data and to perform analysis a modern data platform is required. A cloud hosted data and analytics platform will be built to deliver organisation wide managed self-service analytics, operational reporting and new data driven capabilities including descriptive and predictive analytics. Support for Power BI, python and R will be included. Data governance and security will be incorporated by design. Real-time feeds out to agencies may be enabled in the future.

A high-level view of a modern data platform is below. Architecture governance of the data platform will sit within the overall enterprise architecture function and considers how data is provisioned and provided.

The platform will include data governance by design and will incorporate data policies, frameworks, and standards. Tools for data lineage, metadata and cataloguing will be included, as will identity access and management, data classification and retention policies.

As this will be a cloud-based platform, cloud governance will also be required.

Automated pipelines will ingest data from the various internal and external sources.

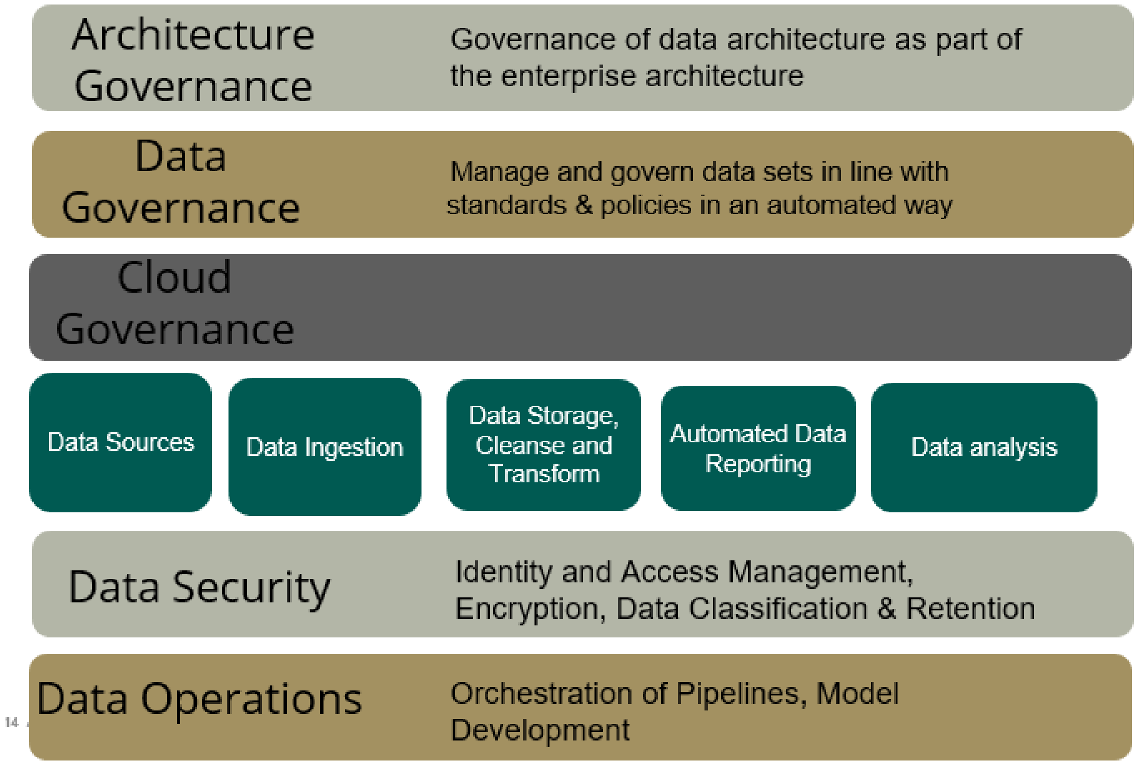
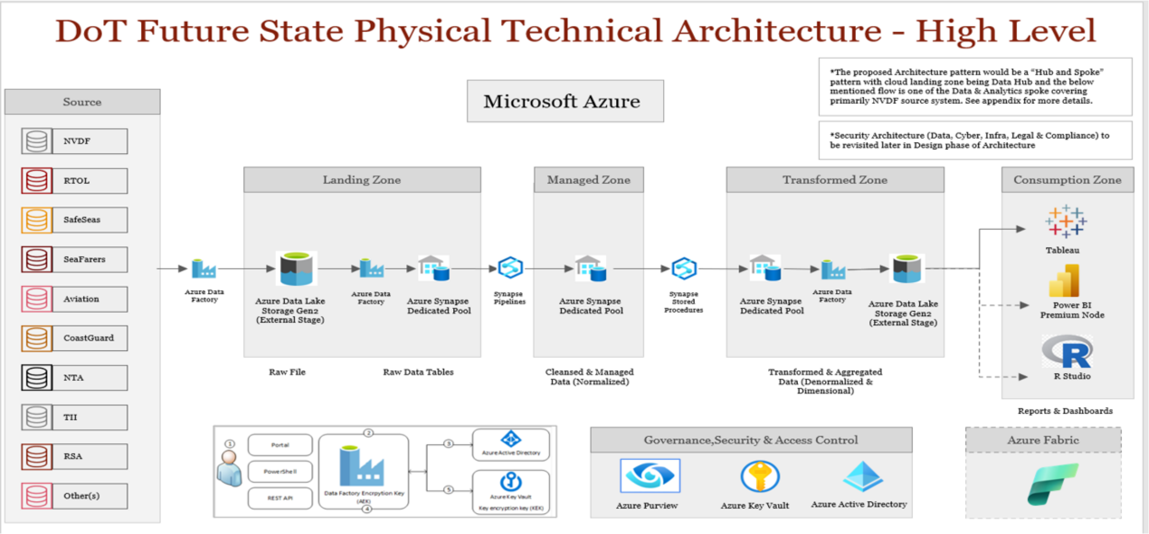

Architecture Governance
Governance of data architecture as part of the enterprise architecture
Data Governance
Manage and govern data sets in line with standards & policies in an automated way
Cloud Governance
Data Sources
Data Ingestion
Data Storage, Cleanse and Transform
Automated Data Reporting
Data analysis
Data Security
Identity and Access Management, Encryption, Data Classification & Retention
Data Operations
Orchestration of Pipelines, Model Development
14

Figure 5 Modern Data Platform Architecture



*The proposed Architecture pattern would be a "Hub and Spoke" pattern with cloud landing zone being Data Hub and the below mentioned flow is one of the Data & Analytics spoke covering primarily NVDF source system. See appendix for more details.

*Security Architecture (Data, Cyber, Infra, Legal & Compliance) to be revisited later in Design phase of Architecture

Figure 7 Technical View of Architecture

Governance

Governance will be needed to enable control and transparency of data and data projects and to enable good decision making and efficiency. The overall governance structure will fall into three categories – strategic, managerial, and operational.

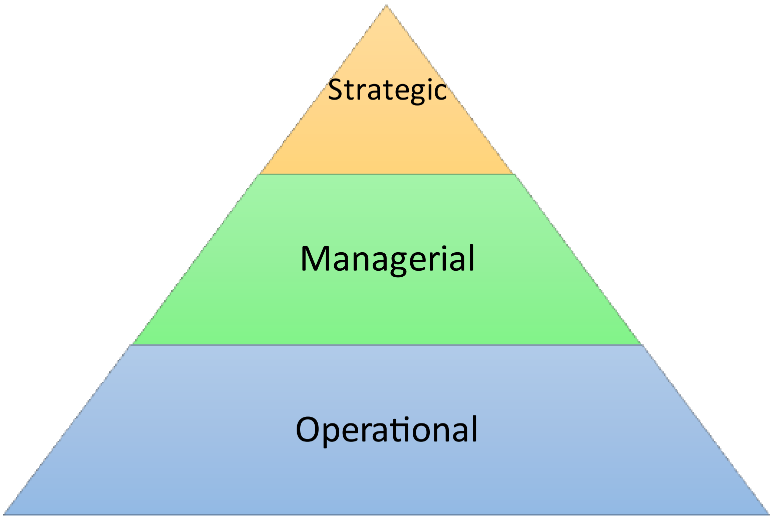


Figure 8 Governance Model

DR

Strategic governance will include agreeing and monitoring the execution of the data strategy and ensuring its alignment to the Department strategy, oversight of technology aspects, budget and resource planning, approval of new projects and ongoing review of DAAD's service provision.

The prioritisation of new projects must align with the overall strategic aims of the department and give consideration to the resourcing of the Data and Analytics team. The framework also needs to factor in ad hoc and urgent requests that DAAD will be required to service. **Error! Reference source not found.** on the following page gives an example processes diagram for assessing new data projects.

Managerial governance will involve the oversight and monitoring of project delivery, manage, and approve change requests to existing projects and stakeholder engagement.

Operational governance will cover the ongoing management of data and compliance management.

A Data Governance and Projects Board will be set up, headed by the DAAD PO. At Strategic level, the Data Board will report into the Department's Technology Steering Board. At Operational level, the Data Board will manage and maintain the quality, security, and accessibility of data in the department. The Board will oversee the governance of data sharing between the Department, its agencies, other PSBs and third parties.

Staff with responsibilities for data management and data compliance in the following areas will report into the Data Governance Board:

a) Data quality
b) Security and Data Protection
c) Data sharing
d) Data Discoverability
e) Metadata

DATA QUALITY

Data quality must be considered from the point data enters the Department through to when it is consumed by the end user. There are multiple aspects to data quality including standardisation, missing data and erroneous data. Quality by design should be integrated where possible.

Data Ingress to the Department

Data enters the Department's system from multiple sources in multiple ways. For example, when new cars are purchased or cars are imported, information from Vehicle Registration Tax is added to the NVDF. Similarly, data from the eCOC (electronic Certificate of Conformity) is also added. Motor tax customers pay either online or at a local motor tax office and these data make their way into the NVDF. Data on road haulage operator licences is gathered in the RTOL system. This will require business owners to be assigned at each stage of data ingress. Each business owner will have responsibility for incoming data quality for their assigned data. They will need to proactively engage with the team or organisation responsible for the source data. They will liaise with the downstream data consumers and will report to the Data Governance Board on issues of data quality under their remit.

Data Ingress to the Data and Analytics Platform

Further quality checks need to be built into the automated process that loads data into the various tiered stages of the central repository. In this flow, standards will have to be applied and missing data dealt with appropriately. Outliers and implausible value will need to be flagged

DRAFT

and corrected if needed. The definition of the standards and rules and the appropriate level of quality will be owned by a senior member of DAAD who will report into the Data Governance and Projects Board and will involve input from multiple stakeholders.

UNIQUE IDENTIFIERS

Where appropriate and useful, unique identifiers should be added to datasets to allow deeper insights to be gained from linkage to other datasets both internally and externally. The key unique identifiers for the department will be the Eircode and the PPSN. Wherever address is captured, Eircode should also be captured. The design of any new application or web interfaces that are required to capture address should ensure Eircode is front and centre. Internally Geocoding tools should be run against datasets that have addresses to increase Eircode coverage. Consideration should be given to the use of PPSN as the unique identifier for the person where appropriate, including where there is a strong business case, e.g., the need to verify identity in order to use a service.

SECURITY AND DATA PROTECTION

Data must always be protected from unauthorised access. Identity management systems must restrict and monitor access to personal data. The storage of personal data on the analytics platform must be minimised and anonymisation and pseudonymisation techniques must be used on ingress to the platform.

Data Processing

Data on the analytics platform needs to be classified based on the privacy level. Privacy level must vary from fully personal data to anonymised and pseudonymised data to data that are not considered private. A system needs to be developed jointly by the data protection officer and DAAD which allows staff who need to access data on the analytics platform to apply for appropriate access to data based on its privacy level. A DPIA should be in place to cover project/team-based processing. If access is granted following the application procedure it should be for a period of one year and if continued access is required, the application must be renewed. The DPO needs to keep records on who has access to what data.

DATA SHARING

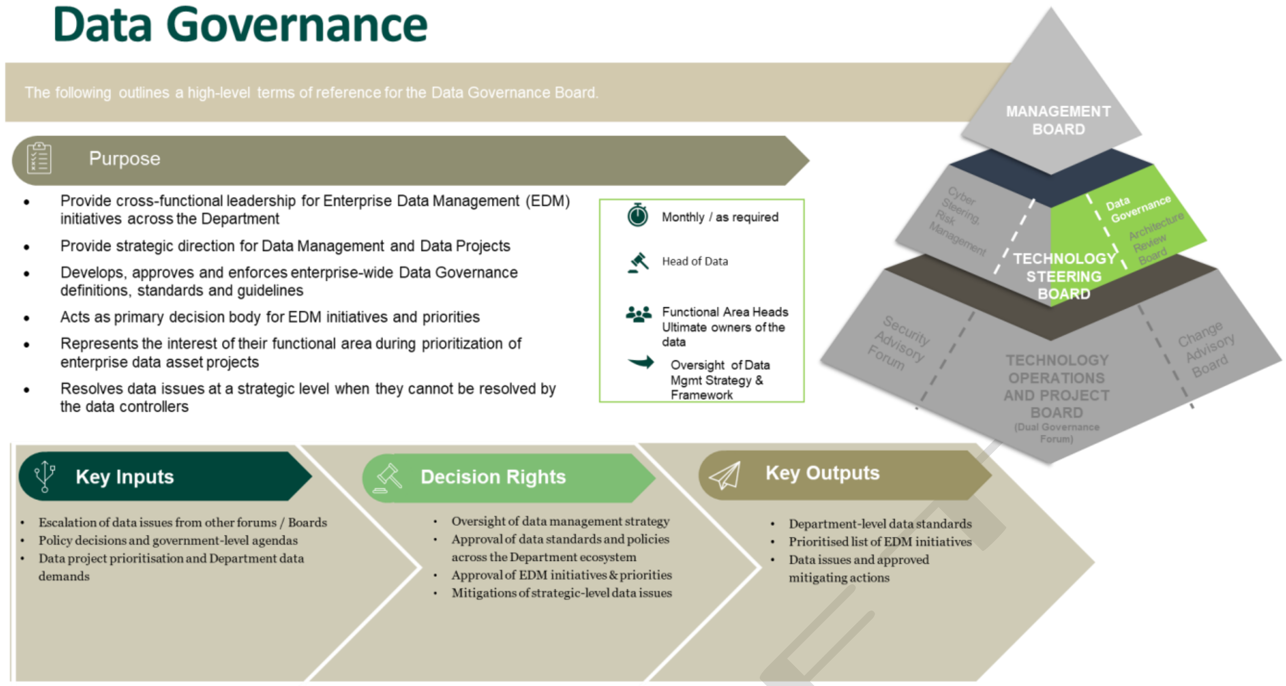
The Department's data officer (DO) must work to encourage safe and appropriate data sharing between the Department, its agencies and other public service bodies. This strategy will align with Principal 4 of the Public Service Data Strategic – "Data is reusable".

"Implementing the Once-Only principle, by ensuring that citizens and businesses supply the same information only once and reducing the number of independent copies of data held in the Public

Service, will promote sharing and reuse of data and common services to improve service provision and decision making."

DATA DISCOVERABILITY AND METADATA

This strategy has proposed a federated model and so data users may not be experts in the data platform and in the depth and breadth of available data. Therefore, discoverability and usability of data is key. All data in the analytics platform needs to have high quality metadata to describe it. The platform should also have a search function to allow data to be easily found. Metadata management should be built into the technology rather than being maintained manually in data dictionaries.



Data Governance

The following outlines a high-level terms of reference for the Data Governance Board.

Purpose

- Provide cross-functional leadership for Enterprise Data Management (EDM) initiatives across the Department
- Provide strategic direction for Data Management and Data Projects
- Develops, approves and enforces enterprise-wide Data Governance definitions, standards and guidelines
- Acts as primary decision body for EDM initiatives and priorities
- Represents the interest of their functional area during prioritization of enterprise data asset projects
- Resolves data issues at a strategic level when they cannot be resolved by the data controllers

Monthly / as required

Head of Data

Functional Area Heads Ultimate owners of the data

Oversight of Data Mgmt Strategy & Framework

Key Inputs

- Escalation of data issues from other forums / Boards
- Policy decisions and government-level agendas
- Data project prioritisation and Department data demands

Decision Rights

- Oversight of data management strategy
- Approval of data standards and policies across the Department ecosystem
- Approval of EDM initiatives & priorities
- Mitigations of strategic-level data issues

Key Outputs

- Department-level data standards
- Prioritised list of EDM initiatives
- Data issues and approved mitigating actions

MANAGEMENT BOARD
Cyber Steering, Risk Management
Data Governance
Architecture Review Board
TECHNOLOGY STEERING BOARD
Security Advisory Forum
TECHNOLOGY OPERATIONS AND PROJECT BOARD
(Dual Governance Forum)
Change Advisory Board

Figure 9 Governance Structure

DRAFT

Operating Model

The operating model presented in this chapter sets out the responsibilities for various aspects of the data and analytics lifecycle.

Federated Model for Analytics and BI

The federated model proposed in this strategy involves the centralised governance and management of data with decentralised data users.

DAAD Overall Responsibilities

DAAD will have the following responsibilities.

- Specifying acceptable data quality standards
- Centrally managing and governing data
- Overall responsibility for Data Sharing Agreements
- Ingesting Data from Agencies and feeding reports out to divisions
- Enabling divisions to leverage insights from data dependant on their data maturity – descriptive, diagnostic, prescriptive analysis
- Stewardship of data throughout the transport eco-system
- Advocating for data usage and upskilling
- Creating advanced analytic models including machine learning
- Working with CSO on NDI, open data and more

Data Quality

DAAD will have overall responsibility for specifying acceptable quality standards for data in the Department. Data quality needs careful attention at key points as data flows from source to user. The sections below discuss the responsibilities for implementing and monitoring the standards specified by DAAD at various nodes in the data flow.

DATA ENTERING THE DEPARTMENT

As discussed in the previous chapter, responsibility for the quality of data entering the Department's system needs clear ownership. Each significant source of incoming data needs a business owner familiar with the data and its uses. The acceptable standards will be defined by DAAD and the data source business owner will report on data quality from their area of responsibility to a project team within DAAD who will have oversight of data quality in the department.

DATA QUALITY WITHIN THE ANALYTICS PLATFORM

As data is extracted into the analytics platform it will need to be further cleansed and standardised. DAAD will have responsibility for specifying the data standards and quality measures and for communicating same to the technology team implementing the platform.

Data Security

Everyone is responsible for data security. However, clear responsibilities are outlined in the previous chapter. DAAD and the DPO need to agree an approach to minimise the amount of personal data entering the analytics platform. DAAD, the DPO and technology need to collaborate on where best to perform any pseudonymisation – the earlier the better.

Responsibility for the technical aspects of implementing the agreed upon approach will be implemented by the technology partner designing the platform. The more automated and integrated the security and authentication procedures are the more likely they are to be adhered to.

Data Sharing

NVDF DATA

A new procedure has been put in place for requesting access to NVDF data. It is available at the link below.

gov.ie - Motor Tax and Vehicle Ownership (NVDF) (www.gov.ie)

Maintenance of Data Sharing Agreements

DAAD will be responsible for the maintenance of DSAs and MoUs. Currently all data sharing that the Department engages in is covered by either primary or secondary legislation. MoUs should be in place for each data sharing setup to set out the legal basis for the sharing of personal data, to highlight the security measures put in place and to specify the parties' expectations of each other with respect to data protection obligations. DAAD review all data sharing agreements and MoUs and take ownership of same.

Where data sharing agreements are in place for non-personal data, the option of open data should be considered instead.

Technology

The technology aspect of the data architecture will be the responsibility of DAAD via contractors. Governance of the data architecture will fall to the enterprise architecture function.

Data and the Department's Agencies

The Department should play a central stewardship role in ensuring that the data collected across the agencies is sufficient to allow detailed policy analysis. This role should include ensuring the agencies collect the right data and adhere to the appropriate standards. Where possible data should be collected once and then shared as needed.

When the Departmental central repository is built there is an opportunity for NTA and TII to share aggregated data (statistics) with the Department efficiently. NTA could share public transport and micromobility usership data, while TII could share aggregated road data and cycling data (they are working on a national approach to collecting this). DAAD will have the ability to combine the aggregated data with our own fleet and driver data to gain detailed insights for policy creation and evaluation.

Furthermore, there are opportunities for efficiencies in how data are currently exchanged between the agencies and the Department. Currently there are numerous transfers of data from NTA to different divisions in the Department. DAAD and the NTA have agreed that once the DAAD platform is built it would make more sense for DAAD to take in the required data from NTA and distribute the required reports to the divisions internally. This would also be an opportunity to streamline and rationalise. The central management of these reports will have the advantage of allowing comparability between the various pieces of analysis which they feed as it can be ensured that reports are built from a consistent base. This is also true for other agencies.

This will however be a change for the divisions concerned and communication will be key.

DAAD and other Divisions in the Department

DAAD will be a centre of excellence for data and analytics in the Department. It will be responsible for upskilling and enabling divisions to use data effectively for policy evaluation and creation.

Communicating this strategy to Divisions will also be important, both for raising awareness and for preparing for some changes that are discussed in the previous section.

DAAD will also work to increase the data skill level throughout the Department. It is hoped and expected that Divisions will be increasingly build up their statistical and analytic capability among their staff, with support from DAAD as outlined above. This will in turn allow DAAD to focus its resources on more advanced analytics including machine learning, AI, predictive modelling etc

Linkages with CSO and Open Data

The Department currently sends aggregated data on new vehicle registrations, penalty point offenses and learner permits to the CSO both to comply with Eurostat regulations and for publishing as Open Data on their PxStat platform. The CSO are currently developing a new PxStat page specifically for DoT where we can take responsibility for uploading our own open data. This page will have DoT branding and the contact details for queries will be DoT staff. Work is currently underway to increase the amount of Open Data the Department will publish in this way. The focus will be on taking the regular PQ and FOI requests received by the Department and publishing these as Open Data. Over time this should reduce the burden on the Department, with those requesting information being able to obtain this through Open Data in a self-service manner.

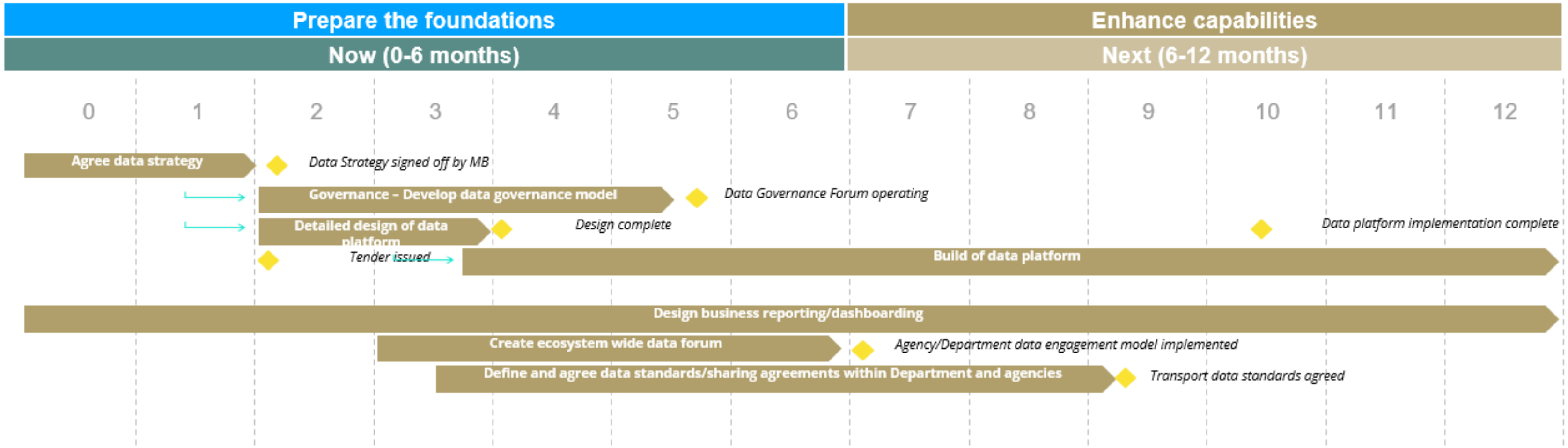
The Department also sends a cut of the NVDF to the CSO on a regular basis. The Department is committed to playing its part to enhance the National Data Infrastructure (NDI), which involves using unique identifiers – Eircode, PPSN and Unique Business Identifier – on key public data sets. Refreshing the Eircode coverage on the NVDF is a key requirement for the Department as specified in earlier sections and this also ensures the Department is contributing to the NDI.

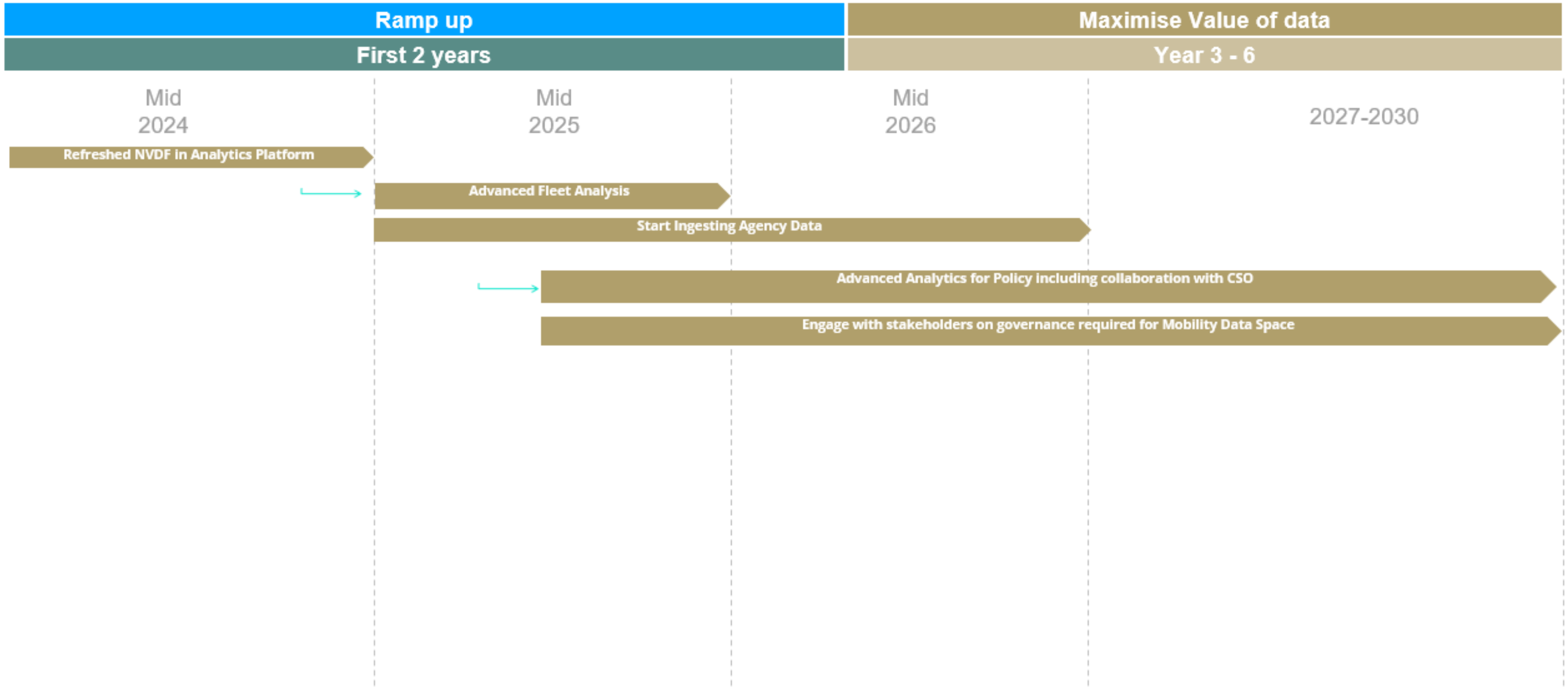
In the future, a lot of the advanced analysis proposed in this strategy will involve linking transport data with CSO data. Collaboration here has use potential to create valuable insight for citizens and policy makers and to create more official statistics for the transport sector.

An Roinn Iompair
Department of Transport

Road Map

Roadmap


Prepare the foundations
Now (0-6 months)
Enhance capabilities
Next (6-12 months)
0
1
2
3
4
5
6
7
8
9
10
11
12
Agree data strategy
Data Strategy signed off by MB
Governance – Develop data governance model
Data Governance Forum operating
Detailed design of data platform
Design complete
Data platform implementation complete
Tender issued
Build of data platform
Design business reporting/dashboarding
Create ecosystem wide data forum
Agency/Department data engagement model implemented
Define and agree data standards/sharing agreements within Department and agencies
Transport data standards agreed



Ramp up

First 2 years

Maximise Value of data

Year 3 - 6

Mid 2024

Mid 2025

Mid 2026

2027-2030

Refreshed NVDF in Analytics Platform

Advanced Fleet Analysis

Start Ingesting Agency Data

Advanced Analytics for Policy including collaboration with CSO

Engage with stakeholders on governance required for Mobility Data Space

↳ Dependencies ◆ Milestones

An Roinn Iompair
Department of Transport